



Department of Computer Science
& Information Engineering

資 訊 工 程 系

人工智慧與邊緣運算實務

03 資料集建置與標註

雲端計算 (Cloud Computing)

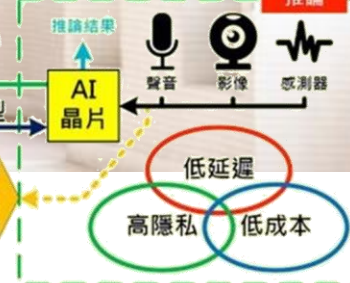
訓練 / 推論 / 儲存



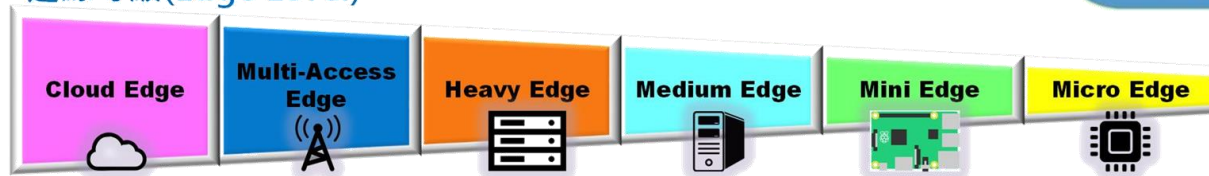
雲端伺服器
Cloud Server

邊緣計算 (Edge Computing)

推論



邊緣等級 (Edge Level)



資訊工程系 許哲豪 助理教授

簡報大綱

➤ 3.1 資料集建置

- 工作流程
- 資料類型
- 資料取得
- 資料擴增

➤ 3.2 公開資料集

- 小型資料集
- Pascal VOC
- ImageNet
- Microsoft COCO
- 其它

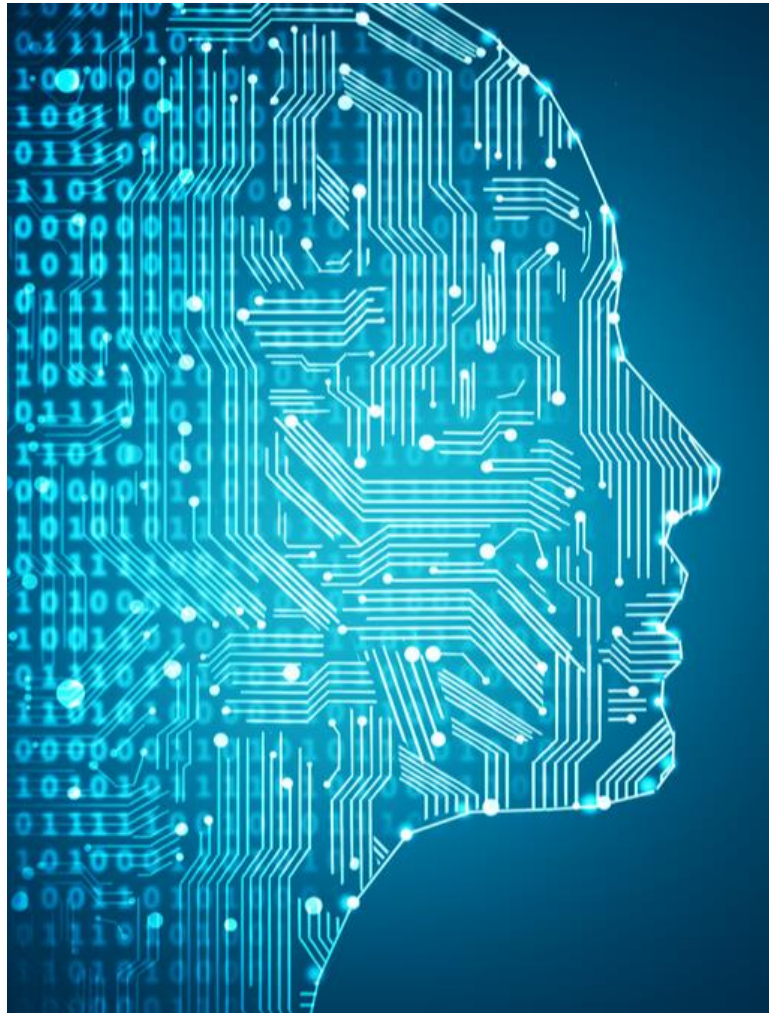
➤ 3.3 資料集標註

- 影像標註類型
- 影像標註格式
- 影像標註工具

➤ 3.4 資料集迷思

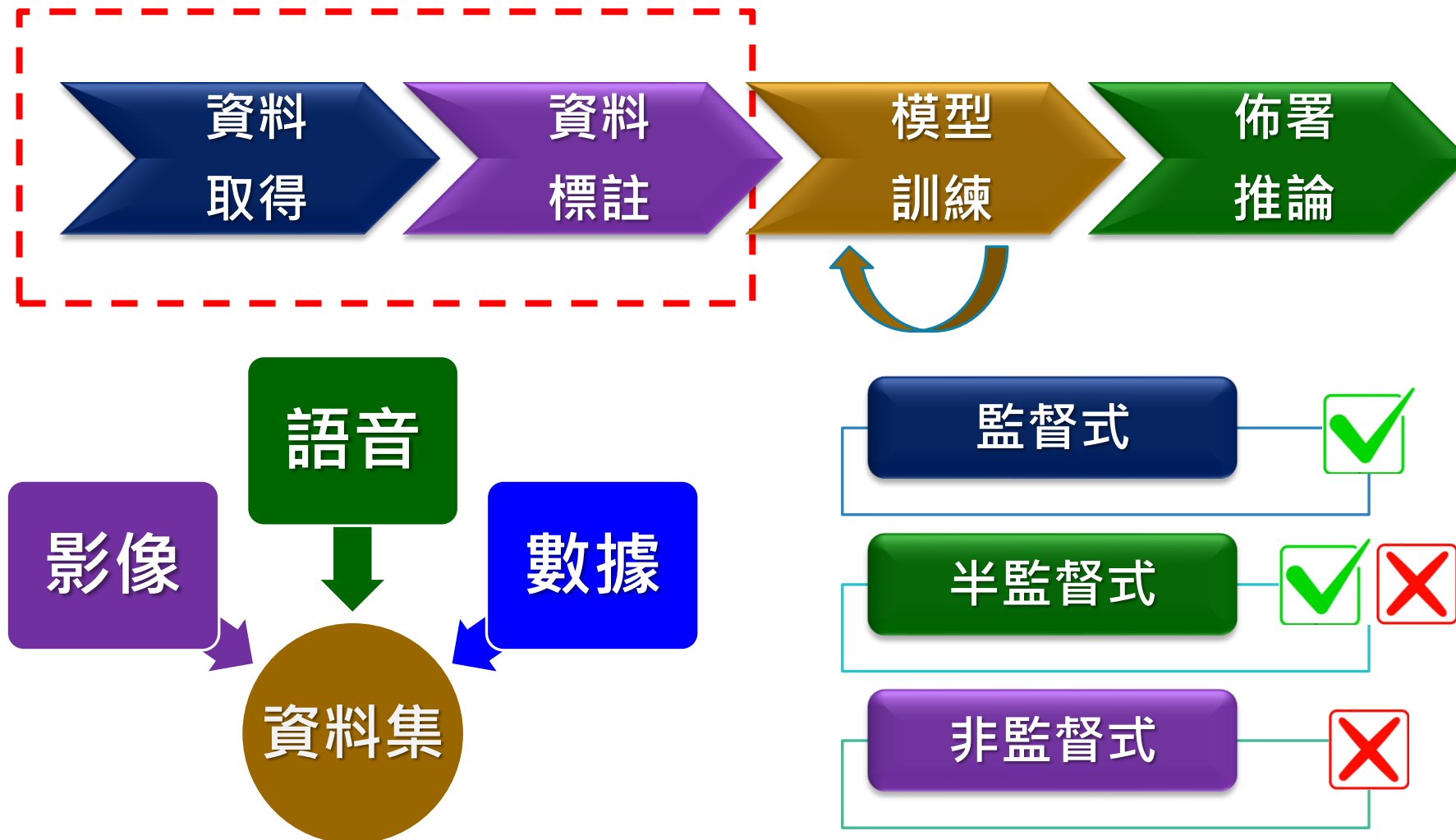
- 資料增長
- 標註水準
- 子集不均
- 自動聚類

3.1 資料集建置



- 工作流程
- 資料類型
- 資料取得
- 資料擴增

3.1.1 工作流程



3.1.2 資料類型—應用領域

➤ 電腦視覺(影像、影片)

- 影像分類
- 物件偵測
- 語義 / 實例分割
- 文字辨識(OCR)
- 生物辨別(人臉、指紋...)
- 圖像標題
- 影像生成、合成
- 行為(骨架)分析
- 軌跡預測
- 品質檢測(瑕疵、分級...)

➤ 自然語言

- 語音、文字互換
- 語音客服、助理
- 語義(文章)理解
- 語言翻譯

➤ 數據分析

- 時序預測
- 生產優化
- 回歸分析
- 推薦系統
- 資訊安全

3.1.2 資料類型—資料型態

➤ 格式化資料

- 一維時序資料 (感測器)
光電、溫度、濕度、壓力、
聲音、速度、運動...
- 二維 / 三維影像資料
可見光、紅外光、深度影像
- 影像序列資料 (影片)
原始、壓縮、串流
- 多維表單數據 (資料庫)

➤ 非格式化資料

- 自然語言、文章
- 影像、影片內容
- 非定時定量收集的資料
- 偶發性資料

3.1.3 資料取得

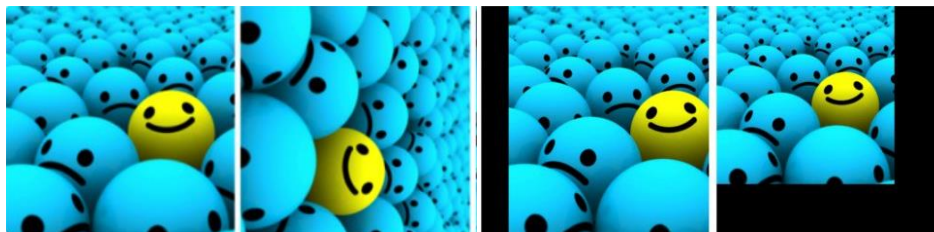
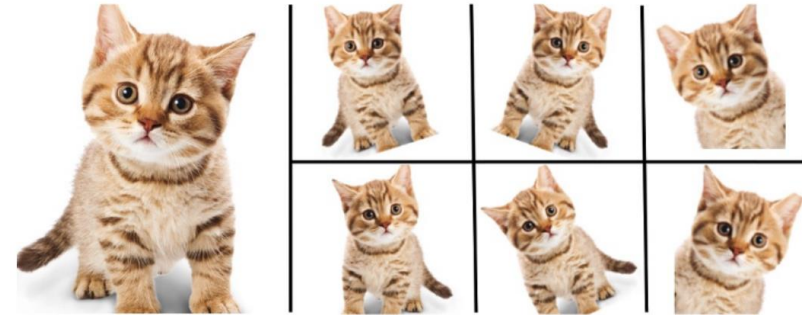
➤ 資料集來源

- 公開資料集
- 私有資料庫
- 網路收集
- 自行拍攝取像
 - 專業相機 (手機、相機、攝影機)
 - 網路攝影機 (USB)
 - 開發板專用相機 (CSI)
 - 可見光
 - 紅外線
- 感測器資料
 - 網路連接、MQTT

➤ 取像要領

- 取像品質近似
 - 解析度、模糊度...
- 多視點
 - 遠近、角度、物件正反面...
- 多光源差異
 - 環境亮度、陰影、光圈、快門...
- 多背景、品種
- **建立足夠多樣性的資料集**
- **資料清洗**
 - 去冗、填空、修補、抑雜訊

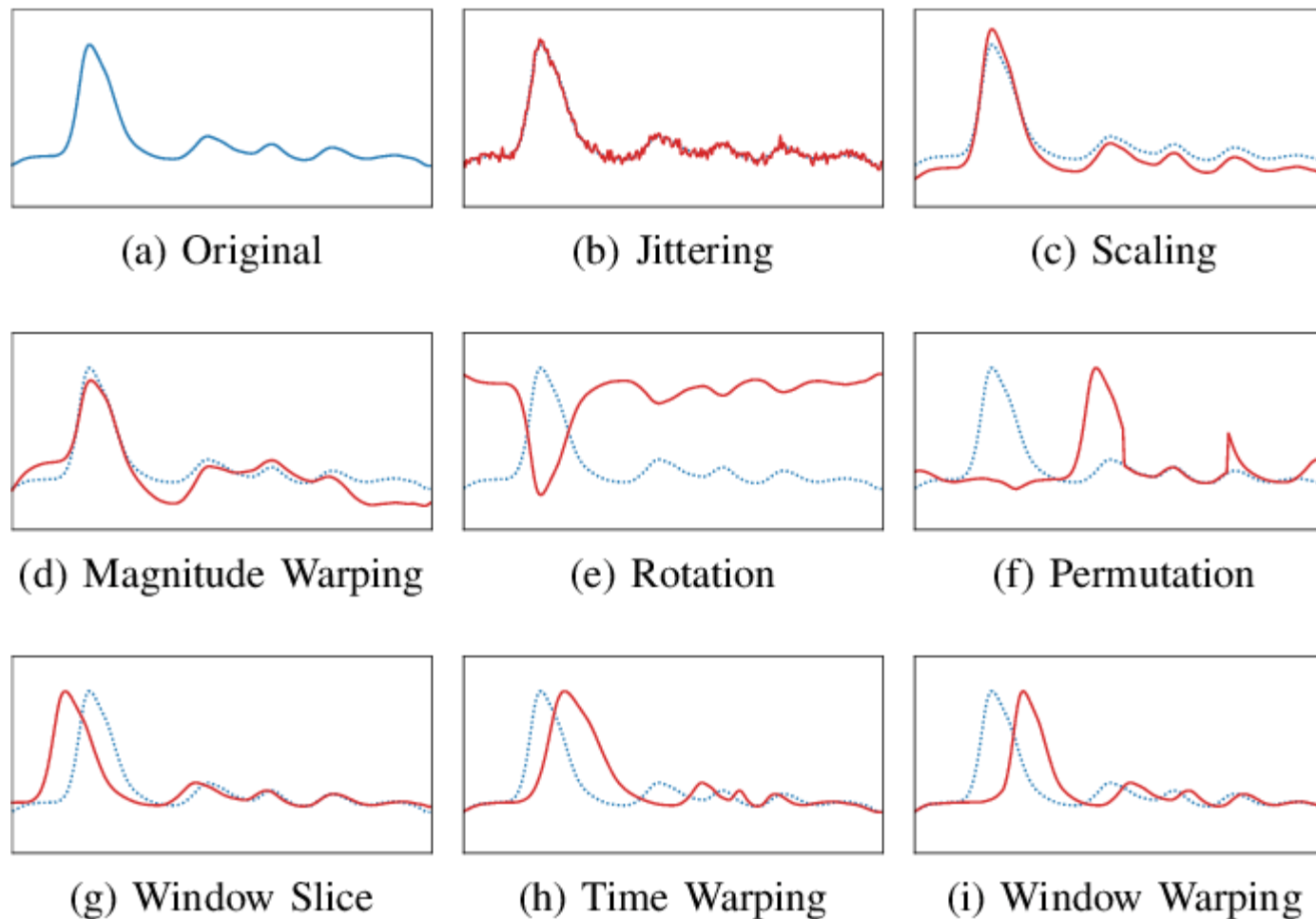
3.1.4 資料擴增—影像類



隨機 平移、縮放、翻轉、鏡射、旋轉、截切、仿射、亮度、對比、色彩、色調、飽和度、模糊、銳利...

資料來源：<https://nanonets.com/blog/data-augmentation-how-to-use-deep-learning-when-you-have-limited-data-part-2/>

3.1.4 資料擴增—時序類



感測器類

光電
溫度
濕度
壓力
聲音
速度
運動
...

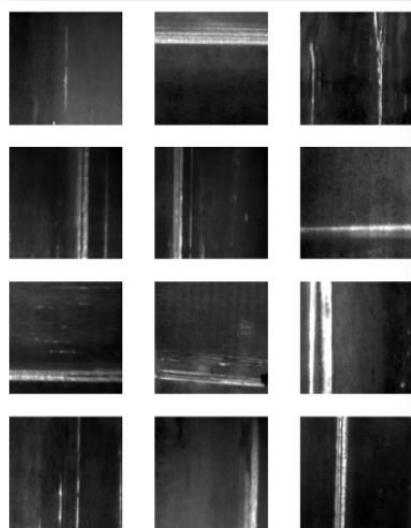
資料來源：<https://arxiv.org/abs/2004.08780>

3.1.4 資料擴增—生成類

資料不易取得、案例數量懸殊、多樣性不足、難以標註

生成方式

虛擬場景生成
生成對抗網路
雜訊添加破壞
影像局部合成



工業表面瑕疵

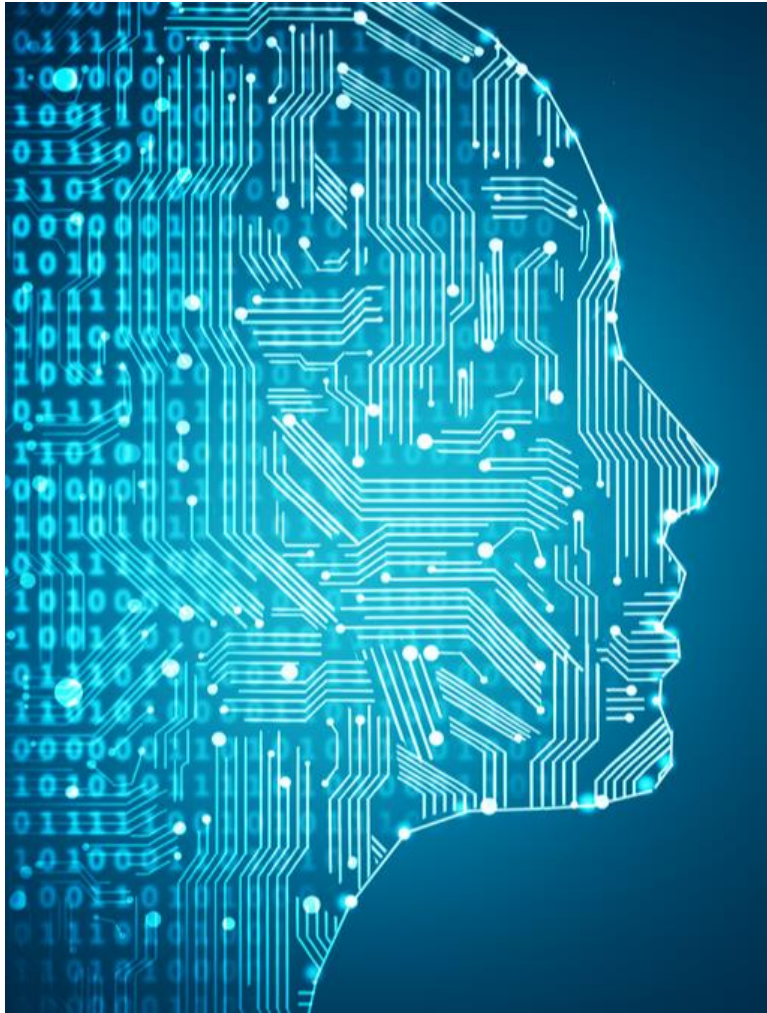
布里斯托大便分類表		
第一類		粒狀、硬身
第二類		腸狀、起塊
第三類		腸狀、表面有裂紋
第四類		長條狀、光滑而柔軟
第五類		一坨坨、但有清晰分界、柔軟
第六類		鬆軟小塊、呈糊狀
第七類		流質、沒有粒塊

大便健康分類



街景分割

3.2 公開資料集

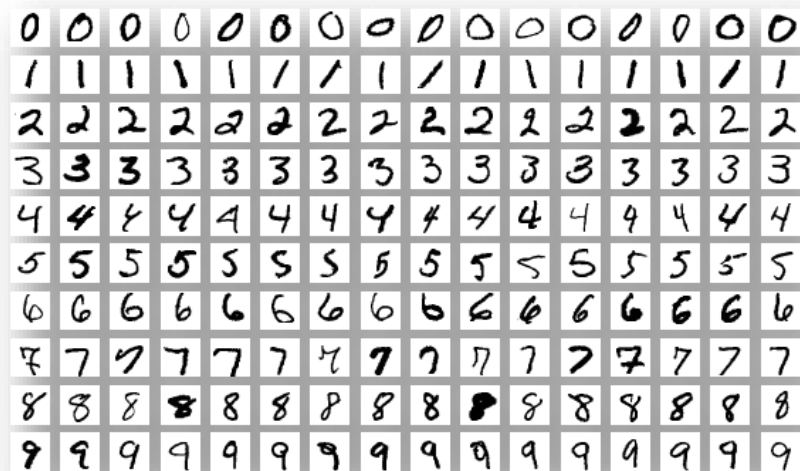


- 小型資料集
- Pascal VOC
- ImageNet
- Microsoft COCO
- 其它資料集

3.2.1 小型資料集

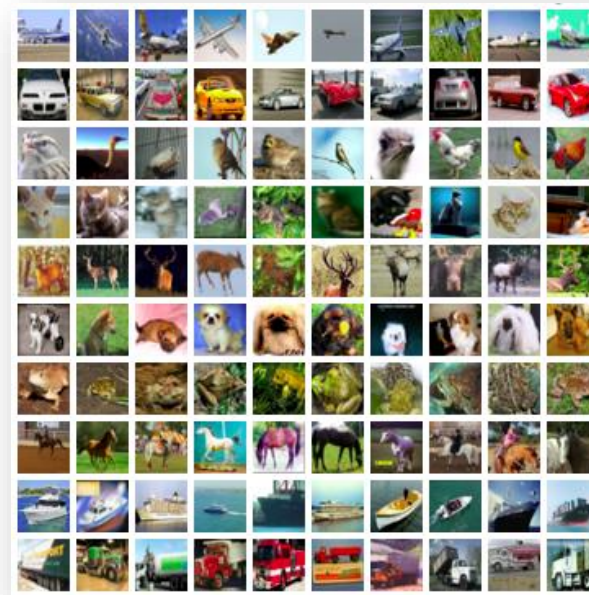
➤ MNIST 手寫數字

- 28*28灰階影像
- 共分十類，數字 0 ~ 9
- 6萬張訓練集，1萬張測試集



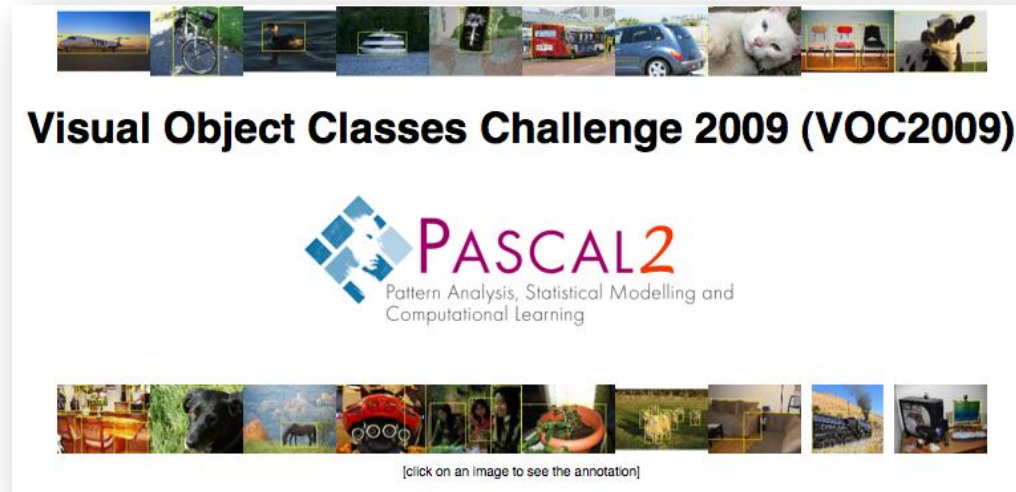
➤ CIFAR-10 彩色影像

- 32*32 彩色影像
- 共分十類，飛機、貓、狗...
- 6萬張，每類6千張影像



3.2.2 Pascal VOC

Pattern
Analysis
Statical Modeling and
Computational
Learning
Visual
Object
Classes



Pascal VOC (2005 ~ 2012)

共有1萬7千多張影像，分為20類，標註內容包括影像分類、物件偵測、語義分割等。其標註資料主要採用**XML**格式。

<http://host.robots.ox.ac.uk/pascal/VOC/>

3.2.3 ImageNet

ImageNet
Large
Scale
Visual
Recognition
Challenge



ImageNet (2010 ~ 2017)

共有超過1400萬張影像，透過Amazon Mechanical Turk外包協助下進行手動標註，包含2萬多個類別，超過100萬個物件偵測邊界框(Bounding Box)被標註。

<http://image-net.org/>

3.2.4 Microsoft COCO

Common
Objects
In
Context



MS COCO (2015 ~)

共有超過32萬張影像，包含91(80)個類別，超過250萬個物件測邊界框被標註。其標註資料主要採用**JSON**格式。

<https://cocodataset.org/>

3.2.5 其它資料集—資料類型

➤ 影像類

- 人臉、動作、表情辨識
- 物件分類、偵測
- 空拍、街道影像
- 影像分割、深度估測

➤ 影片類

- 色情、暴力、粗俗

➤ 文字類

- 新聞、訊息、問答
- 摘要、知識圖譜

➤ 聲音類

- 演說、音樂、信號
- 機器翻譯、語義理解

➤ 生物醫療

- X光、超音波、感測

➤ 數據類

- 金融、天氣、網路、遊戲
- 時間序列、資訊提取
- 異常偵測、推薦、決策

3.2.5 其它資料集 (1 / 2)



- 國網中心資料集平台 **(32,288)**

<https://scidm.nchc.org.tw/>



- 政府資料開放平台

<https://data.gov.tw/>



- Kaggle(一個數據建模和數據分析競賽平台)

<https://www.kaggle.com/>

3.2.5 其它資料集 (2 / 2)



WIKIPEDIA
The Free Encyclopedia

- Machine-Learning Datasets (38)

https://en.wikipedia.org/wiki/List_of_datasets_for_machine-learning_research



- Paper with Code – datasets (5,630)

<https://paperswithcode.com/datasets>

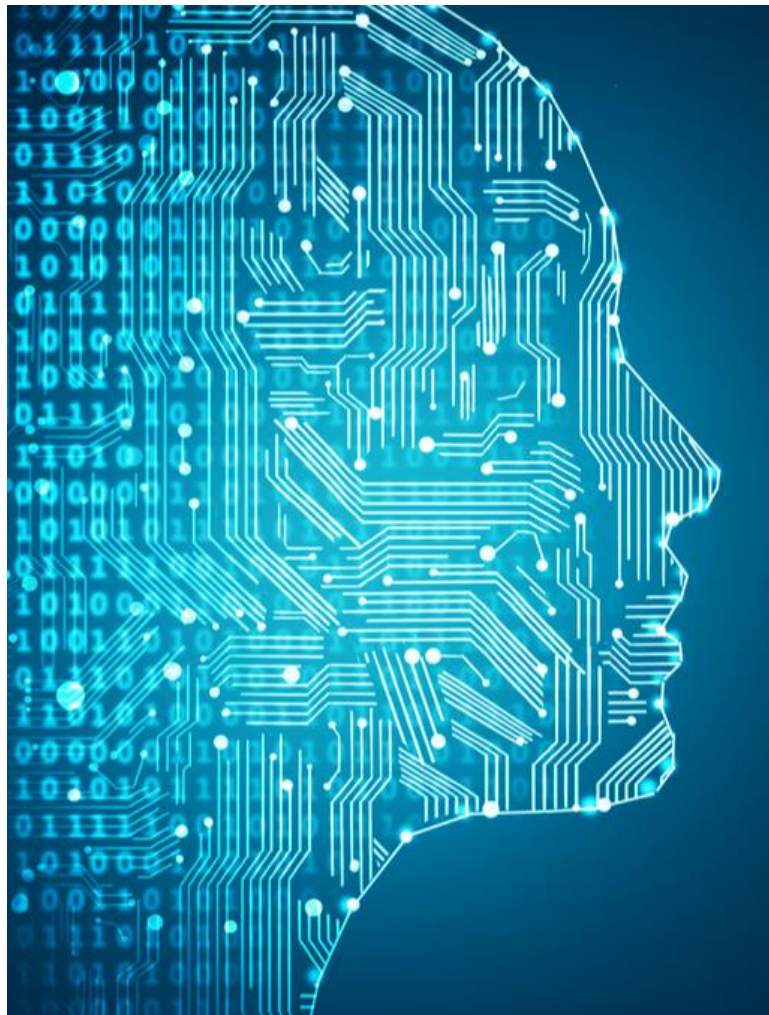
COMPUTER VISION
O.N.L.I.N.E
YOUR GATEWAY TO THE COMPUTER VISION WORLD

- Computer Vision Online – Datasets (119)

<https://computervisiononline.com/datasets>



3.3 資料集標註

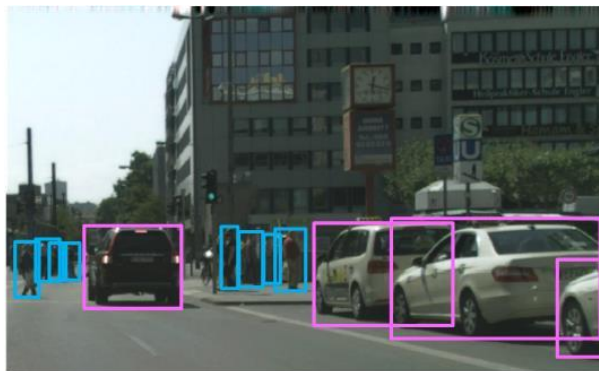


- 影像標註類型
- 影像標註格式
- 影像標註工具

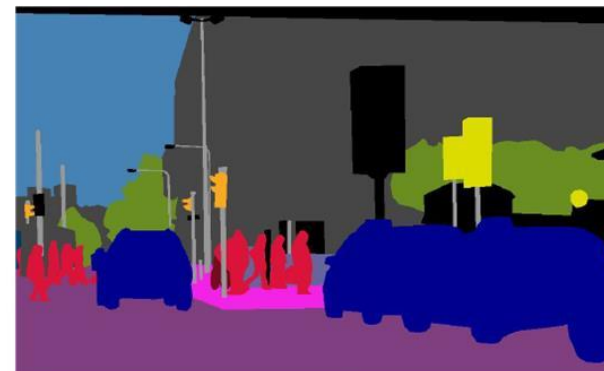
3.3.1 影像標註類型—依工作內容



(a)



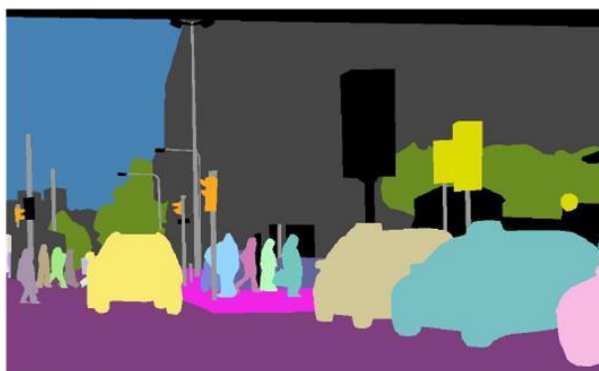
(b)



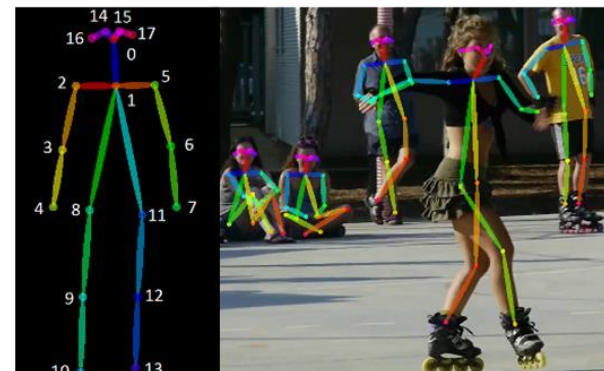
(c)



(d)



(e)



(f)

(a)原始影像/影像分類，(b)物件偵測，(c)語義分割，(d)實例分割，
(e)全景分割，(f)人體姿態（關鍵點）。

3.3.1 影像標註類型—依標註外形



(a)



(b)



(c)



(d)



(e)

(a)點/線，(b)矩形，(c)多邊形/貝茲曲線，(d)自由筆刷，(e)超像素。

3.3.2 影像標註格式—VOC (xml)



```
<annotation>
  <folder>JPEGImages</folder>
  <filename>img_1020.jpg</filename> 影像名稱
  <path>C:\VOC\JPEGImages\img_1020.jpg</path>
  <source>                                影像路徑
    <database>Unknown</database>
  </source>
  <size>
    <width>640</width> 影像尺寸
    <height>480</height>
    <depth>3</depth> 色彩通道數
  </size>
  <segmented>0</segmented>
  <object>
    <name>tomato</name> 物件名稱
    <pose>Unspecified</pose> 物件姿態
    <truncated>0</truncated> 影像是否被截切
    <difficult>0</difficult> 影像是否困難辨識
    <bndbox>
      <xmin>360</xmin> 物件框左
      <ymin>227</ymin> 上角座標
      <xmax>395</xmax> 物件框右
      <ymax>276</ymax> 下角座標
    </bndbox>
  </object>
  <object>
    ... 其它物件資訊
  </object>
</annotation>
```

第一個小蕃茄標註資料

3.3.2 影像標註格式—COCO (json)

COCO 資料格式

```
{
  "info": info,
  "images": [image],
  "annotations": [annotation],
  "licenses": [license],
}

info{
  "year": int,
  "version": str,
  "description": str,
  "contributor": str,
  "url": str,
  "date_created": datetime,
}

image{
  "id": int,
  "width": int,
  "height": int,
  "file_name": str,
  "license": int,
  "flickr_url": str,
```

```
  "coco_url": str,
  "date_captured": datetime,
}

license{
  "id": int,
  "name": str,
  "url": str,
}

annotation{
  "id": int,
  "image_id": int,
  "category_id": int,
  "segmentation": RLE or [polygon],
  "area": float,
  "iscrowd": 0 or 1,
  "bbox": [x,y,width,height],
}

categories[{
  "id": int,
  "name": str,
  "supercategory": str,
}]
```

Segmentation : [polygon]

```
"segmentation":
[[510.66,423.01,...,510.45,423.01]],
"area": 702.1057499999998,
"iscrowd": 0,
```

Segmentation : [RLE]

```
"segmentation": {"counts":
[20736,2,453,5,452,9,447,13,444,...,5,34552],
"size": [457,640]
},
"area": 3074,
"iscrowd": 1,
```

3.3.2 影像標註格式—YOLO (txt)

VOC 物件偵測資料格式

```
<size> 影像尺寸
  <width>353</width> 影像寬度
  <height>500</height> 影像高度
  <depth>3</depth> 色彩深度(通道數)
</size>
<segmented>0</segmented> 是否分割
<object> 物件 (目標) 訊息
  <name>dog</name> 物件名稱 (20分類)
  <pose>Left</pose> 拍攝角度 (前後左右及未定)
  <truncated>1</truncated> 是否被遮擋或截斷
  <difficult>0</difficult> 是否難以檢測
  <bndbox> 物件 (目標) 外框
    <xmin>48</xmin> 物件左上角X座標
    <ymin>240</ymin> 物件左上角Y座標
    <xmax>195</xmax> 物件右下角X座標
    <ymax>371</ymax> 物件右下角Y座標
  </bndbox>
</object>
```

YOLO 物件偵測資料轉換

$$X = (xmin + (xmax - xmin) / 2) * 1.0 / width$$

$$Y = (ymin + (ymax - ymin) / 2) * 1.0 / height$$

$$W = (xmax - xmin) * 1.0 / width$$

$$H = (ymax - ymin) * 1.0 / height$$

邊界框位置
以左上表示

邊界框位置
以中心表示

YOLO 物件標註檔案格式(*.txt)

[分類編號ID] [物件中心X座標] [物件中心Y座標] [物件寬度佔比W] [物件高度佔比H]

0 0.344192634561 0.611000000000 0.416430594901 0.262000000000

OmniXRI Oct. 2020 整理繪製

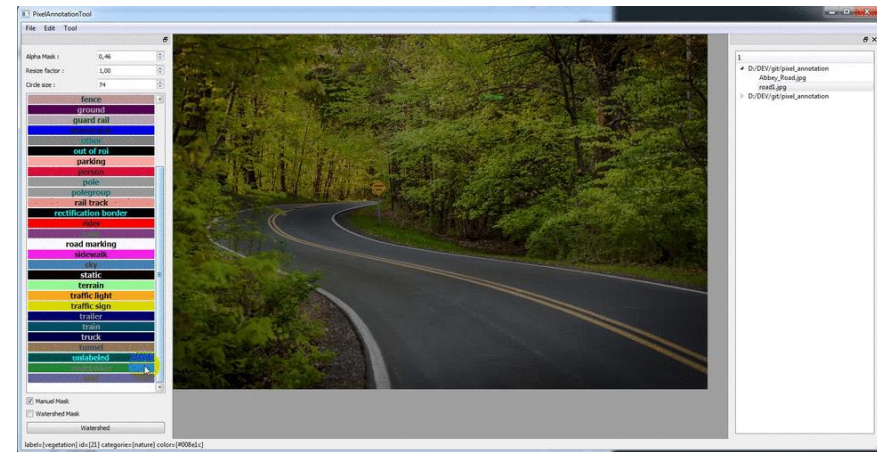
3.3.3 影像標註工具—分類與偵測

➤ 【影像分類/物件偵測類】

- RectLabel <https://rectlabel.com> (商業付費版)
- **Labellmg** <https://github.com/tzutalin/labellmg>
- Labelme <https://github.com/wkentaro/labelme>
- **Intel CVAT** <https://github.com/openvinotoolkit/cvat>
- OpenCV SuperAnnotate Desktop <https://opencv.org/superannotate-desktop> (部份免費)
- labelbox <https://github.com/Labelbox/Labelbox>
- Microsoft VoTT <https://github.com/microsoft/VoTT>
- VGG Image Annotator (VIA)
<http://www.robots.ox.ac.uk/~vgg/software/via>

3.3.3 影像標註工具—影像分割

➤ 【影像分割類】



➤ PixelAnnotationTool

<https://github.com/abreheret/PixelAnnotationTool>

➤ superpixels-segmentation

<https://github.com/Labelbox/superpixels-segmentation>

➤ semantic-segmentation-editor <https://github.com/Hitachi-Automotive-And-Industry-Lab/semantic-segmentation-editor>

3.3.3 影像標註工具

➤ labellmg (本地)

- 開源圖形化介面標註工具
由Python和QT5所開發
- 支援多種平台安裝
(Windows, Linux, Mac)
- 主要提供物件框位置、大小及標籤標註功能
- 提供PASCAL VOC(xml)、YOLO(txt)格式輸出

➤ Intel CVAT (線上/本地)

- 由INTEL開源可使用Docker安裝或直接線上操作
- 提供邊界框、多邊形、折線及點形式標註
- 具有半自動標註功能適合連續影片快速標註
- 支援多種框架輸入輸出(VOC, YOLO, TFrecored, MOT, ImageNet...)

線上版：<https://cvat.org>

3.3.3 影像標註工具—安裝labelImg

- GITHUB <https://github.com/tzutalin/labelImg>
- 最簡單安裝方式在Python 3.x環境下使用pip安裝（會自動安裝PyQt5, lxml）

pip3 install labelImg

- 啟動標註程式
labelImg

（其它作業系統安裝方式
請參閱GITHUB 說明）



labelImg執行畫面

3.3.3 影像標註工具—下載範例

➤ 至Github下載為ZIP至桌面並解壓縮

https://github.com/OmniXRI/20201024_AIGO_Lab2

—20201024_AIGO_Lab2

└ Ex1_Annotations (影像標註練習檔)

└ VOC2007

└ Annotations (影像標註結果 *.xml)

└ ImageSets (影像集清單)

└ Main (訓練、驗證、測試清單)

└ JPEGImages (原始影像 *.jpg)

或依上述路徑安排方式自行準備相關影像檔案

3.3.3 影像標註工具—設定路徑



3.3.3 影像標註工具—邊界框標註



3.3.3 影像標註工具—CVAT簡介

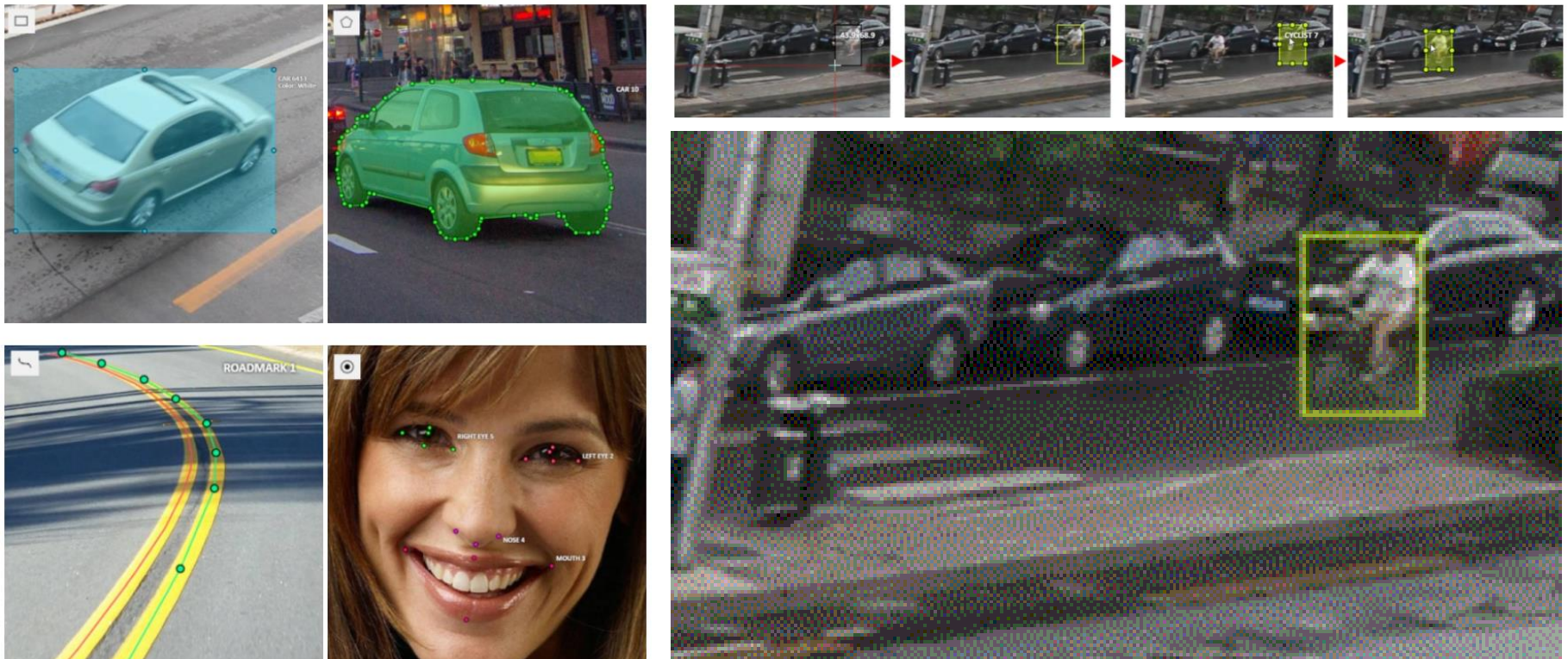
➤ Intel Computer Vision Annotation Tool (CVAT)

- 支援多種標註格式輸出入(CVAT, VOC, YOLO, COCO, Tfrecorecord, MOT, LabelMe ...)
- 可輸出 PNG格式影像多分類分割遮罩(mask)檔
- 可線上標註或以Docker於本地端安裝伺服器
- 支援影像分類、物件偵測、影像分割等任務。
- 標註方式：點、折線、方框、多邊形。
- 支援影片自動標註（補間追蹤特定物件）功能，節省標註時間。
- 僅適用Google Chrome瀏覽器。

更完整說明：<https://github.com/openvinotoolkit/cvat>

3.3.3 影像標註工具—線上CVAT

- 線上版標註服務(<https://cvat.org/>)免安裝，建立新帳戶後即可直接使用（每個帳戶最多10個任務，上傳內容 < 500MB）



3.3.3 影像標註工具—CVAT (Youtube)



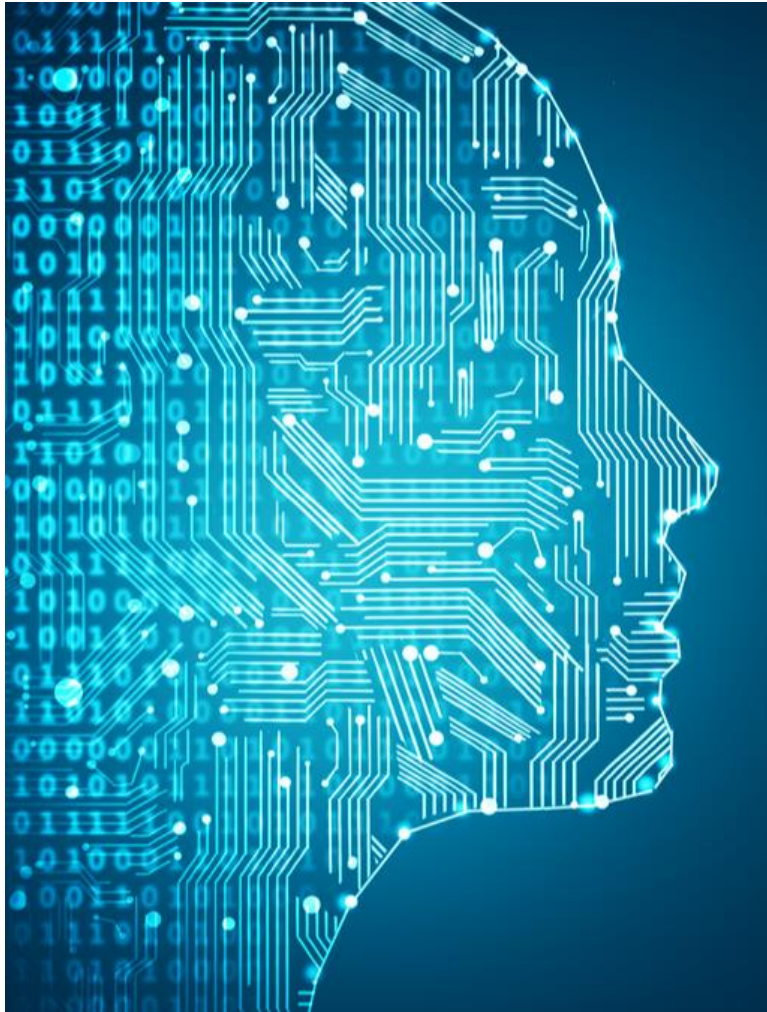
➤ CVAT Computer Vision Annotation Tool | OpenVINO™ toolkit | Ep. 51 | Intel Software

<https://www.youtube.com/watch?v=BKLyHOEACFw>

➤ CVAT Auto Annotation | OpenVINO™ toolkit | Ep. 52 | Intel Software

<https://www.youtube.com/watch?v=jbqOa8DX7Jg>

3.4 資料集迷思

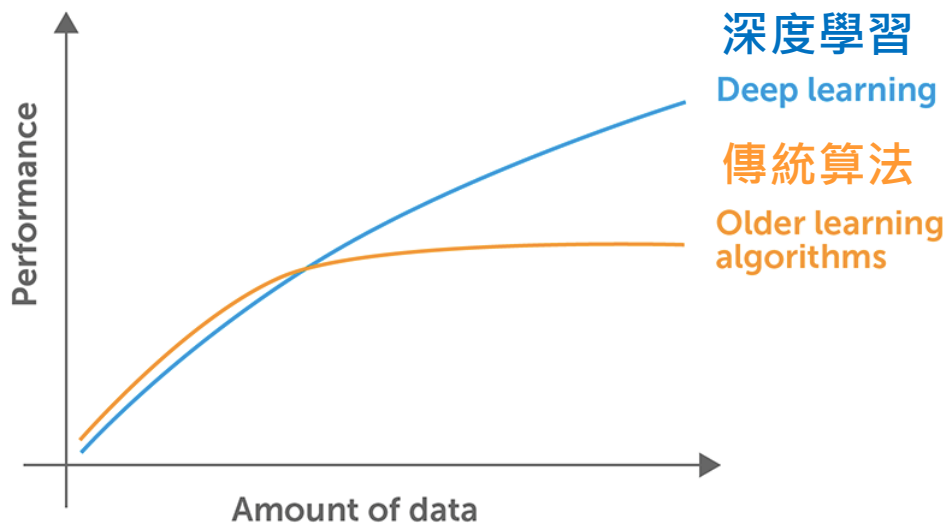


- 資料增長
- 標註水準
- 子集不均
- 自動聚類

3.4.1 資料增長

* AI 會自我學習只要一直提供資料就會變得更厲害？

資料增長 vs. 訓練成果



圖片來源：吳恩達

- 擴增資料集、反覆訓練
 - 監督式 (分類)
 - 非監督式 (聚類)
 - 遷移式 (監督型加速式)
- 免資料集、明確獎懲規則
 - 增強式 (遊戲、閉路控制)
- 少量資料集、反覆訓練
 - 生成對抗式 (GAN)

不想學習圖書館再多書也沒用

3.4.2 標註水準

* 數據越多學習訓練效果越好？

➤ 監督式

➤ 標註品質

- 標註工具（便捷、管理）
- 普通人 vs. 專家

➤ 資料多樣性

- 資料重新採集
- 資料衍生擴增
- 收集公開資料集
- 對抗生成

➤ 子類平均度

➤ 非監督式

➤ 增強式—獎懲規則驅動免資料集（如遊戲）

➤ 關連式無標註型資料集越多越好（如語音、翻譯）

➤ 高維資料聚類

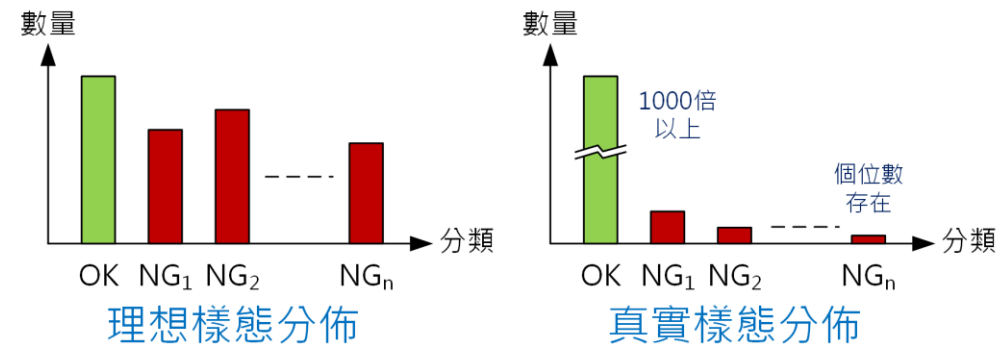
- 格式化資料及樣態種類
- 資料降維、特徵提取
- 計算複雜度（運算速度）
- 記憶體容量

高品質標註資料越多越好

資料在精不在多

3.4.3 子集不均

* 只要大量良品就能自動學習特徵做為不良品檢測？



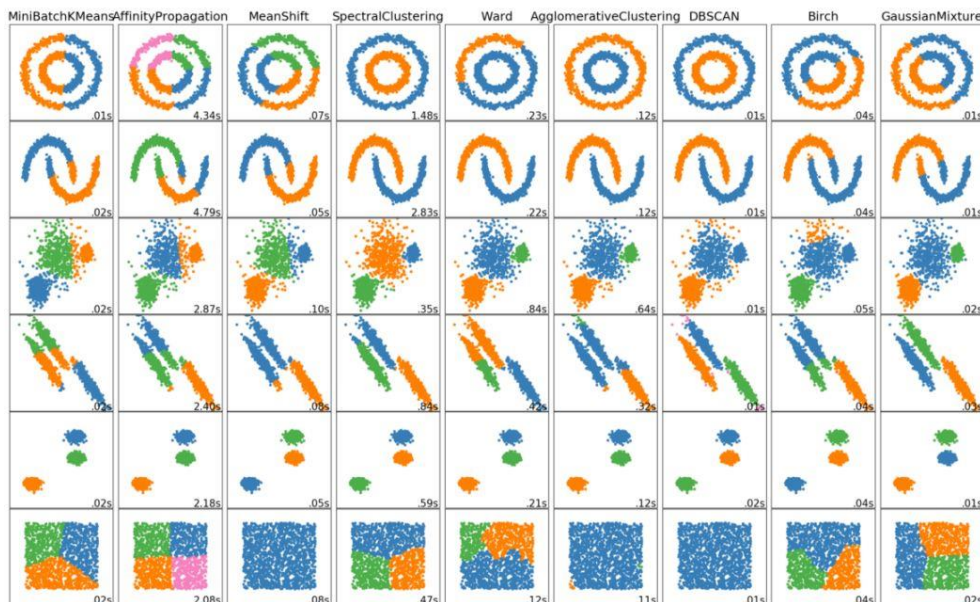
- 人類採特徵法學習只須少量樣本，深度學習採暴力法（填鴨死記）須巨量樣本。
- 異常偵測法 (Anomaly Detection) 採資料統計法找出合理分佈區間，再挑出離群樣本，適用資料維度較小的的案例。

自動提取特徵，工程耗大

3.4.4 自動聚類

* 提供巨量未標註資料就能自動找出分類規則關連？

常見聚類算法及分類結果



(二維資料分類)

圖片來源：<https://www.plob.org/article/12370.html>

➤ 翻譯問題

➤ 表面上為非監督式學習，但實質上訓練素材（語料）為人類產生之標註資料。

➤ 影像（語音）分類問題

➤ 高維（超稀疏）資料自動聚類困難

➤ 半監督式

➤ 部份有標註，部份自動產生標註，再由人決定標註正確性，加快標註時效。

困難重重，有待突破

小結

➤ 資料集建置

- 依據需求使用公開資料集或自行收集相關資料集，並保持多樣性及進行初步整理清洗，不足處再加以擴增。

➤ 公開資料集

- 可從多種管道中取得不同型態已標註之資料集，電腦視覺、自然語言、格式數據皆有對應資料集可供參考。

➤ 資料集標註

- 了解資料標註方式，資料存放格式及基本工具操作。

➤ 資料集迷思

- 資料在精不在多，標註品質決定訓練成果，子集不均需調整，自動聚類仍不成熟。

參考文獻

- 許哲豪，”【AI HUB專欄】如何建立精準標註的電腦視覺資料集”

https://omnixri.blogspot.com/2020/10/ai-hub_16.html

- 許哲豪，”【課程簡報分享】AI萬能？導入AI的八大迷思剖析”

<https://omnixri.blogspot.com/2019/08/aiai.html>